

C	7	596.8	17.9	936	14	B0954456	B0954456	ACN	COUR
C	8	593.8	17.8	653	10	AM994023	AM994023	RC3	BN003
C	9	595.8	17.5	609	12	BE966851	BE966851	ACN	BN003
C	10	593.6	16.9	879	12	BF979362	BF979362	602	288452
C	11	559.2	16.8	608	14	BM790880	BM790880	K	ESTF0070
C	12	558.4	16.7	1096	12	BG245770	BG245770	602	358561
C	13	552	16.6	569	9	AI301140	AI301140	901	6005
C	14	546.4	16.4	848	13	BI562019	BI562019	603	325069
C	15	539.2	16.2	744	13	BI526944	BI526944	602	9271124
C	16	539.2	16.2	1008	12	BF979911	BF979911	602	288452
C	17	538.8	16.2	731	13	BG964392	BG964392	602	832010
C	18	536.4	16.1	559	14	B0083862	B0083862	K	ESTF0146
C	19	534	16.0	779	13	BI692921	BI692921	603	342760
C	20	527.2	15.8	849	12	BF581951	BF581951	602	099402
C	21	525.8	15.8	530	9	AI672868	AI672868	w672904	x
C	22	525.2	15.8	530	9	AL708563	AL708563	DKF2P686C	
C	23	521.8	15.7	774	13	BI828492	BI828492	603	078270
C	24	518.2	15.5	533	10	BE218431	BE218431	hV386L0	x
C	25	517.8	15.5	552	10	AM137691	AM137691	UI-H-B11-	
C	26	515.6	15.5	797	12	BF133383	BF133383	601	759179
C	27	509	15.3	521	12	BF531260	BF531260	744c08	x
C	28	507.6	15.2	529	10	BE081676	BE081676	QV2	-BT063
C	29	503.8	15.1	611	10	BB618558	BB618558	BB618558	
C	30	498.2	14.9	555	10	AM966212	AM966212	EST378285	
C	31	496.6	14.9	616	10	BB618535	BB618535	UI-H-B11-	
C	32	495	14.8	516	14	BQ021096	BQ021096	UI-H-DH1-	
C	33	492.8	14.8	507	12	BE769819	BE769819	RC2	-PT004
C	34	490.8	14.7	906	12	BF141940	BF141940	601	790430
C	35	488.4	14.6	528	12	BF744833	BF744833	QV2	-BT063
C	36	485.6	14.6	493	14	BM855444	BM855444	K	ESTF0138
C	37	485	14.5	892	12	BF535659	BF535659	602	054056
C	38	484.4	14.5	528	9	AA890726	AA890726	aK13008	s
C	39	481	14.4	493	9	AI889507	AI889507	w005d10	x
C	40	478.8	14.4	494	9	AI565607	AI565607	tq79c03	x
C	41	474.4	14.2	862	13	BI081766	BI081766	602	879683
C	42	469.4	14.1	484	9	AI683933	AI683933	tW54h11	x
C	43	468.8	14.0	497	9	AA399077	AA399077	zE93h12	x
C	44	466.6	14.0	810	12	BF164971	BF164971	601	777513
C	45	462.4	13.9	477	9	AI923722	AI923722	w059h10	x



Db 661 ACAGACAATGCTACTGACCTCCGATCCGATGATGAGGCCGCTCTCTAACTCCGTC 720

QY 2260 ACAGCTACGCTCTGCGAGGAGTATCATCCCGCAGAGGAGGACGCTCAAGGCCGG 2319

Db 721 ACAGCTACGCTCTGCGAGGAGTATCATCTCTCTGAGGATGGCTACAAAGCCGG 780

QY 2320 GTTTCCTCAGAGCTGTCAGGCTGTCGATCGACTTATGGGGGACAGACGATGACT 2379

Db 781 GTGCTCTACAGCTGTCAGGCTGTCGAGCTGACTTGTGGAGTGC -ACAGACATGACC 839

QY 2380 CTGATGGAATACACCTCCAGCTGGAACCTTTCCAGCTGGAGTCTGGGAATACGGAC 2439

Db 840 TTGAGAGA---TCGTTCCCAAGAACTTTTCCAGAGACTC---GGGAATCCCGAC 892

QY 2440 GTGATCTCTTTTATAGTCCAAATGATGTGACCCAGTCTGCACTGCTGGAGATCAAC 2499

Db 893 ATATCTCTCTTT-TAGATCCAAATGATGTGACCCAGTCTGCACTTCT -GAGATCAAC 950

QY 2500 ACCATCCGCGCTCAGATGCAATCCAAATGAAAGCTGCTCTGGACCCCTGAGTACCCAGC 2559

Db 951 ACCATCCGCGCTCAGATGCAATCCAAATGAAAGCTGCTCTGGACCCCTGAGTACCCAGC 1010

QY 2560 ACAGCTCAGATGAGGAGCTGATGAGTGTGCAACTTCTCTGTCGAGAGCGCGCT 2619

Db 1011 ATGAGCTCTGATAGAGCTGATGAGTGTGCAACTTCTCTGTCGAGAGTGTGCGC 1070

QY 2620 GCTTGCCCGCTCTGCTAGTGGCTGACTACATGCTATGCTGACGACCTGTCGTCGCGG 2679

Db 1071 GCTTGCTCAGCTCTGCTGCGCTCTGACTTACACACTTTTGTGACGACCTGTGTCGCGG 1130

QY 2680 ATCCAGAGAGACTTACGTCAGTGTGCGAGAAACCCAGCTATGCTCTGTCGATTTCTCT 2739

Db 1131 ATCCAGAGAGACTTACGTCAGTGTGCGAGAAACCCAGCTATGCTCTGTCGATTTCTCT 1190

QY 2740 CCTAGAGAGAGAGTACCAATCTGCAAAACCAATGATTTCTGCTGAAAGTGGCCTCT 2799

Db 1191 CCTAGAGAGAGAGTACCAATCTGCAAAACCAATGATTTCTGCTGAAAGTGGCCTCTG 1250

QY 2800 GCAGGACAGCTGCTACTGCTCTGCTCAGCGTCTGACCTGCTGCTGCTGCTGCTGCTGCT 2859

Db 1251 GCTGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1310

QY 2860 CAAAGACTAGATGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2919

Db 1311 CAAAGACTAGATGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1370

QY 2920 CTGCGACAGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2979

Db 1371 CTGCGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1430

QY 2980 TTTACGAGCAAGAA-TGACTCTTTGGGAGATCAATCATTTACCTCCAGAGAGACTCT 3038

Db 1431 TTTACGAGCAAGAGTCTCTCTTTGGGAGATCAATCATTTACCTCCAGAGAGACTCTCT 1490

QY 3039 GATGATTTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3098

Db 1491 GATGATTTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550

QY 3099 GAG--GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3156

Db 1551 GAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1610

QY 3157 GCGATTTGGGAGCGAG--ATGCTGCAACAGCCAGCTGCTGCTGCTGCTGCTGCTGCTG 3215

Db 1611 GCGATTTGGGAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1670

QY 3216 TATCAGATTTGTTGATTTGAGA--TCTTTTATAGATGATGCTGCTGCTGCTGCTGCT 3272

Db 1671 TATCAGAGATTTGATTTGAGACCTGCTTTTATAGAGTACCCAAACCTGCTCTAG 1730

QY 3273 CTTCCTCAAACTGCCAAATATACCA 3300

Db 1731 CTTCCTCAAACTGCCAAATATACCA 1758

RESULT 2  
BQ278843/c 1069 bp mRNA linear EST 07-MAY-2002  
LOCUS BQ278843  
DEFINITION AGENCOURT\_7049684 NIH\_MGC\_107 Homo sapiens cDNA clone IMAGE:5805500  
5', mRNA sequence.  
ACCESSION BQ278843  
VERSION BQ278843.1 GI:20489051  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC <http://mhc.nci.nih.gov/>.  
1 (bases 1 to 1069)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LILCH2044 row: m column: 21  
High quality sequence stop: 690.  
Location/Qualifiers  
1..1069  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5805500"  
/clone\_lib="NIH\_MGC\_107"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: pORF7; Site:1: EcORI;  
Site:2: XhoI; cDNA made by oligo-dT priming,  
directionally cloned into EcORI/XhoI sites using the  
following 5' adaptor: GGCACAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library."  
BASE COUNT 266 a 255 c 292 g 255 t 1 others  
ORIGIN  
Query Match 25.2%; Score 840.6; DB 14; Length 1069;  
Best Local Similarity 92.8%; Pred. No. 8.2e-184;  
Matches 982; Conservative 0; Mismatches 60; Indels 16; Gaps 9;  
QY 2280 ACTCATATCCCGCAGAGTGACAGAGTACAAAGCCGGGTTTCTCT---CACAGCCTG 2335  
Db 1058 ATTCCCGCCCGCAGAGGGGGCAACGCTACAAAGCCGGGTTTCTCTCACAGCCTGTT 999  
QY 2336 TCAGCTTCTCTATGACTT-ATTGGGGTGACAA-CAGATATGACTCTGATGATCAC 2393  
Db 998 CAGGCTTCTCTATGACTTAATTGGGGTGACAAACAGATATGACTTCTGATGGAAATTC 939  
QY 2394 CTCCCGACGTG---ACCTTTCACAGTGG--AGTCTTGGGAAATACCGAGCTGATCTTC 2448  
Db 938 ACTCCCGCAAGTTGAAACTTTTCCACTGCGAAGTCTTGGAAATACCGGAGCTGATCTTC 879  
QY 2449 TTTTATAGT-CCAATGATGACCCAGTCTGCACTTCTGGAGATCAACAC--ATC 2505  
Db 878 TTTTATAGTCCCAATGATGACCCAGTCTGCACTTCTGGAGATCAACACCCATTC 819  
QY 2506 CCGCTCAGGTGAGTGCACAGAAACTGT-CCTTGAAGTTTCTGCTGCCAGAGACTG 2564  
Db 818 GCGCTCAGGTGAGTGCACAGAAACTGTCCCTTGAAGTTTCTGCTGCCAGAGACTG 759  
QY 2565 CTCAGATGAGACTGTGATGCTGCACTTCTCTGTCGAGAGAGCGGCGTGTG 2624

Accession	Source	Organism	Reference Authors	Title	Journal	Comment
Db 758	CTCAGATGGAGACCTGTGTATGGCTGCACATTTCACCTTCTGTGGGAGACGGGGCTGCTTG					
Qy 2625	CCGGCTCTGCTCAGTGGGCTGACACTACCATGCTATGCTGACGACTGTGTGGCTGGGATCCA					
Db 698	CCCGCTGTGCTCAGTGGGCTGACACTACCATGCTATGCTGACGACTGTGTGGCTGGGATCCA					
Qy 2685	GAAGACTACTTACGTGTGTGGCAGAAACCCAACTATGCTCTGTGTGGCAATTCTCTGCTCGA					
Db 638	GAAGACTACTTACGTGTGTGGCAGAAACCCAACTATGCTCTGTGTGGCAATTCTCTGCTCGA					
Qy 2745	GCAGAGAGTGCACATCTGTGCACAAACCATATGATTCTTGCGGTGAAGAGGGGCACTCTGACAG					
Db 578	GCAGAGAGTGCACATCTGTGCACAAACCATATGATTCTTGCGGTGAAGAGGGGCACTCTGACAG					
Qy 2805	CACCTGTACTGTCACATCTGCTACACGCTTGCTGACCTGCTACTTTTGGAAAAAGAAATCAAAA					
Db 518	CACCTGTACTGTCACATCTGCTACACGCTTGCTGACCTGCTACTTTTGGAAAAAGAAATCAAAA					
Qy 2865	ACTAGAGTACAAAGTACTCCAAAGCTGTGTATGATATGCTACTCTCAAGAGACTGTGACCTGCC					
Db 458	ACTAGAGTACAAAGTACTCCAAAGCTGTGTATGATATGCTACTCTCAAGAGACTGTGACCTGCC					
Qy 2925	AGCAGCTACACAGCTGCGGCATCATATGGAAGGAGAGATGTAGAGAGCAGCATCATCTTTAC					
Db 398	AGCAGCTACACAGCTGCGGCATCATATGGAAGGAGAGATGTAGAGAGCAGCATCATCTTTAC					
Qy 2985	CAGCAAGAA-TCACTCTTTTGGGGAAGATCAAAATCATTTTACCTCCAGAGAGACTCCTGATGG					
Db 338	CAGCAAGAAAGTCACTCTTTTGGGGAAGATCAAAATCATTTTACCTCCAGAGAGACTCCTGATGG					
Qy 3044	ATTGATACATCATGCGCGCTGGAAGACATCTCTCAAGAGAGCCCAACATAGACATGACATGTAGAGGC					
Db 278	ATTGATACATCATGCGCGCTGGAAGACATCTCTCAAGAGAGCCCAACATAGACATGACATGTAGAGGC					
Qy 3104	ACTGCTCTGCTACACCTGCTGCTCCTCACTTGAATAGACACTTTGCAAGGCTCGGCGGATTT					
Db 218	ACTGCTCTGCTACACCTGCTGCTCCTCACTTGAATAGACACTTTGCAAGGCTCGGCGGATTT					
Qy 3164	GGGTGCGCAGCATCTCTGCAACACCCACTGCTGGAATCTCTTCATTTGTGGCCTTATCAGAT					
Db 158	GGGTGCGCAGCATCTCTGCAACACCCACTGCTGGAATCTCTTCATTTGTGGCCTTATCAGAT					
Qy 3224	GTTTGAATTTTCAGATCTTTTATATAGATACCCAAACCCCTCTTCTGTGCTTGCCTCAAA					
Db 98	GTTTGAATTTTCAGATCTTTTATATAGATACCCAAACCCCTCTTCTGTGCTTGCCTCAAA					
Qy 3284	CCTGCCAATATATACCCACACTTGTGTTAAATTTAAA 3321					
Db 38	CCTGCCAATATATACCCACACTTGTGTTAAATTTAAA 1					
RESULT 3						
BM830016						
LOCUS	BM830016	705 bp	mRNA	linear	EST 06-MAR-2002	
DEFINITION	K-EST0103421 S21SN0520s1 Homo sapiens cDNA clone S21SN0520s1-17-E12					
ACCESSION	BM830016					
VERSION	BM830016.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	1 (bases 1 to 705)					
	Kim, N. S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.,					
	Oh, K. U., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and					
	Kim, Y. S.					
	21C Frontlier Korean EST Project 2001					
	Unpublished (2002)					
	Contact: Kim YS					
	Genome Research Center					
	Korea Research Institute of Bioscience & Biotechnology					

52, Eoosun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@email.kribb.re.kr  
Plate: 17 row: E column: 12  
High quality sequence stop: 705.  
Location/Qualifiers

### Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S21SN0520s1-17-E12"
/clone_1b="S21SN0520s1"
/sex="F"
/tissue_type="Stomach"
/cell_type="floating aggregates"
/cell_line="SNU-520"

```

/home-0rgn: Stomach: Vector: PT218RP1 Site:1: EcoRI; Site:2: NotI. The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deacapped with tobacco acid pyrophosphatase (TAP). The deacapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okazama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli TOP10<sup>®</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(CT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing cDNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli TOP10<sup>®</sup> with

BASE COUNT	157 a	208 c	169 g	171 t
ORIGIN				

Query Match	21.1%;	Score 705;	DB 14;	Length 705;
Best Local Similarity	100.0%;	Pred. No. 2e-152;		
Matches 705; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY 1960 GTCCAGGCCCTGTGTGCCCTGTGGTCCAGGACCAAGAACAAACAAGATCCACTCTGTGTC 2019

Db I GTCCAGGCCCTGTGTCCTTGTTGGTCCAGGGACCAAGAACACAAATCACCACATCCTCATGTC 80

Db 61 TACATGTTGGACCTTCTCAGCAGACACTCCAGCAGACTTCAACTACACTTCC 120

QY 2080 GCTTTGGCAACACACCGTCACTCTTCTGGAGGCCAACCTTCACTTCCAAGGTTGAAA 2139

Db 121 GCATTGGCAACACCGCTCAGCTCTTGCTGGAGGGCCAAAGCTTCAGTTCCAAAGGGTTCAAA 180

Db 181 TACTTCCATCATCTTTACCCCTAGATCTCTGTGGAACCCAGGCTAGCAAAATCTCTGTGTGC 240

QY 2200 ACCGACAATGTCACGTGACCTCCGGATTCTCTGAGGGTGAGTCAGGGTTCTCCAAATCTATC 2259

Db 241 ACCGACAATGTCTACTGACCTCCGATTCCTGAGGGTGAAGTCAAGGTTCTCCAAATCTATC 300



KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.  
TITLE 21C Frontiers Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 20 row: A column: 12  
High quality sequence stop: 689.  
Location/Qualifiers  
1. 689  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="S21SNU520-20-A12"  
/clone\_lib="S21SNU520"  
/sex="F"  
/tissue\_type="Stomach"  
/cell\_line="Floating aggregates"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ19RP1; Site:1: EcoRI;  
Site:2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tabacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

BASE COUNT 151 a 211 c 154 g 173 t  
ORIGIN

Query Match 20.0%; Score 666; DB 14; Length 689;  
Best Local Similarity 99.7%; Pred. No. 2.1e-143;  
Matches 688; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 1802 ATGGCGTGGCCCTACTAGCCGCTCCCTGAGAACCTCTGATGTGGCTCTCTCT 1861  
|||||  
DB 1 ANGCGGTGGCCCTACTAGCCGCTCCCTGAGAACCTCTGATGTGGCTCTCTCT 60  
|||||  
OY 1862 GCACCTCTTGTCTCTGTTACTATATTGACCGAGATTCCAGAACCTCTCTG 1921  
|||||  
DB 61 GCACCTCTTGTCTCTGTTACTATATTGACCGAGATTCCAGAACCTCTCTG 120  
|||||  
OY 1922 CCCCTAACCAATTCGAAAGCCCAAGCCCTTATGCTGTCAGGCTGTGCTG 1981  
|||||  
DB 121 CCCCTAACCAATTCGAAAGCCCAAGCCCTTATGCTGTCAGGCTGTGCTG 180  
|||||  
OY 1982 GTCCAGGAGCAAGAACCAAGATCACTCTGTGCTACATGATTCACCTTCTC 2041  
|||||  
DB 181 GTCCAGGAGCAAGAACCAAGATCACTCTGTGCTACATGATTCACCTTCTC 240  
|||||  
OY 2042 GCAACACTCCAAACGAGACTTTCACACTCTCTCCGCTTTGGCAACACCGTAC 2101  
|||||  
DB 241 GCAACACTCCAAACGAGACTTTCACACTCTCTCCGCTTTGGCAACACCGTAC 300  
|||||

OY 2102 TTGCTGAGGGCCAGGCTTCACTTCCAAAGGGTGAATACTTCCATCACTTACCCTCA 2161  
|||||  
DB 301 TTGCTGAGGGCCAGGCTTCACTTCCAAAGGGTGAATACTTCCATCACTTACCCTCA 360  
|||||  
OY 2162 GTCTCTGTGGAACCAAGGATAGAAATGTCTGTGTGACCGACAAATGCTACTGACCTCC 2221  
|||||  
DB 361 GTCTCTGTGGAACCAAGGATAGAAATGTCTGTGTGACCGACAAATGCTACTGACCTCC 420  
|||||  
OY 2222 GGATTCCTGAGGGTGTAGTCAGGGTTCTCCAAATCTATACAGCCCTACGCTCCAGGAG 2281  
|||||  
DB 421 GGATTCCTGAGGGTGTAGTCAGGGTTCTCCAAATCTATACAGCCCTACGCTCCAGGAG 480  
|||||  
OY 2282 TCATCATCCCCCAGAGGTGACAGGCTACAA-GGCCGGGGTTTCTCCAGGCTGTACG 2340  
|||||  
DB 481 TCATCATCCCCCAGAGGTGACAGGCTACAAAGGCGGGTTTCTCCAGGCTGTACG 540  
|||||  
OY 2341 CTGTCTGATGCACTTATTTGGGGTGAACAACAGATATGACTGTGATGAAATCACTCCCA 2400  
|||||  
DB 541 CTGTCTGATGCACTTATTTGGGGTGAACAACAGATATGACTGTGATGAAATCACTCCCA 600  
|||||  
OY 2401 GCTGAACCTTTTCCACCTGAGACGCTTGGGAATCCGAGCGATGCTCTTTATAGTCC 2460  
|||||  
DB 601 GCTGAACCTTTTCCACCTGAGACGCTTGGGAATACC-GACGTATCTTCTTTATAGTCC 659  
|||||  
OY 2461 AATGATGTGACCCAGTCTGCTGAGTTCTGGG 2490  
|||||  
DB 660 AATGATGTGACCCAGTCTGCTGAGTTCTGGG 689  
|||||

RESULT 6  
AV751975 685 bp mRNA linear EST 19-OCT-2000  
LOCUS AV751975 NPD Homo sapiens cDNA clone NPDBGF09 5', mRNA sequence.  
DEFINITION AV751975  
VERSION AV751975.1 GI:10909823  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 685)  
AUTHORS Song, H., Peng, Y., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N.,  
Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu,  
M., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Ye, M., Zhang, Q.,  
Han, Z., Chen, Z., Hu, R. and Chen, J.  
Homo sapiens NPD library cDNA clones  
Unpublished (2000)  
CONTACT: Qinghua Zhang  
Shanghai Institute of Endocrinology, Rui-Jin Hospital  
197 Rui-Jin II Road, Shanghai 200025, P. R. China  
Tel: 86-21-64370045(ex. 663332)  
Fax: 86-21-64743206  
Email: mbsliens.stn.sh.cn  
This clone is available at Shanghai Hematology Institute in  
Shanghai.  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.  
Location/Qualifiers  
1. 685  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NPDBGF09"  
/clone\_lib="NPD"  
/tissue\_type="pituitary"  
/dev\_stage="adult"  
/note="Vector: pBluescript sk(-); Site:1: EcoRI; Site:2:  
XhoI"

BASE COUNT 145 a 194 c 173 g 169 t 4 others  
ORIGIN

Query Match 18.8%; Score 625.2; DB 10; Length 685;



Best Local Similarity 96.8%; Pred. No. 5.8e-134;  
Matches 658; Conservative 0; Mismatches 17; Indels 5; Gaps 2;

QY 2043 CAACACCTCCACACCGAGGACTTTCACATCACTTCCCGCTTGGCAACACCGCTCATCT 2102  
DB 1 CAACACTCCACACCGAGGACTTTCACATCACTTCCCGCTTGGCAACACCGCTCATCT 59  
QY 2103 TGTGTGAGAGGCGCAAGCTTCACTTCCAAAGGCTTGAATTAATCTTCATCACTTACCTCCAG 2162  
DB 60 TGTGTGAGAGGCGCAAGCTTCACTTCCAAAGGCTTGAATTAATCTTCATCACTTACCTCCAG 119  
QY 2163 TCTCTGTGGAACACGAGGTAGAAATGTCTGTGTGACCGACAATGTCACTGACCTCCG 2222  
DB 120 TCTCTGTGGAACACGAGGTAGAAATGTCTGTGTGACCGACAATGTCACTGACCTCCG 179  
QY 2223 GATTCCTGAGGCTGATGACAGGCTTCCAAATCATATCATACAGCTTCCAGGAGCT 2282  
DB 180 GATTCCTGAGGCTGATGACAGGCTTCCAAATCATATCATACAGCTTCCAGGAGCT 239  
QY 2283 CATCATCCCCCGACAGGTGACAGGCTCAAGGCGGGCTTCCATCAAGCTGTGACGCT 2342  
DB 240 CATCATCCCCCGACAGGTGACAGGCTCAAGGCGGGCTTCCATCAAGCTGTGACGCT 299  
QY 2343 TGTGTGATGACTTATTTGGGGTGACACAGATATGACTCTGATGGAATCATCTCCGAGC 2402  
DB 300 TGTGTGATGACTTATTTGGGGTGACACAGATATGACTCTGATGGAATCATCTCCGAGC 359  
QY 2403 TGACCTTTTCCACCTGAGAGCTCTTGGGAATACCGGAGCTGATCTCTTTTATAGTCCAA 2462  
DB 360 TGAACCTTTTCCACCTGAGAGCTCTTGGGAATACCGGAGCTGATCTCTTTTATAGTCCAA 419  
QY 2463 TGATGTGACCCAGTCTGACAGTCTTGGGAGATCAACCAACATCCGCGCTGAGGAGTCC 2522  
DB 420 TGATGTGACCCAGTCTGACAGTCTTGGGAGATCAACCAACATCCGCGCTGAGGAGTCC 479  
QY 2523 ACAGAAAACCTGCTCTGGAAGTTTGTCTGTGACAGAACTGTCTCAGATGGAGACTGTGA 2582  
DB 480 ACAG----CTGTCTCTGGAAGTTTGTCTGTGACAGAACTGTCTCAGATGGAGACTGTGA 535  
QY 2583 TGGCTGCAACTTCCACTTCTGCTGTGGGAGAGCGGGGCTGTGCTGCTGCTCACTGAGC 2642  
DB 536 TGGCTGCAACTTCCACTTCTGCTGTGGGAGAGCGGGGCTGTGCTGCTGCTCACTGAGC 595  
QY 2643 TGACTACCATGCTATCTGCTGACAGACTGTGTGGGTGGGATCCAGAAAGACTTACGTTG 2702  
DB 596 TGACTACCATGCTATCTGCTGACAGACTGTGTGGGTGGGATCCAGAAAGACTTACGTTG 655  
QY 2703 GCGAGAACCCAGCTATGCT 2722  
DB 656 GCGAGCCGCTATGCTGTGT 675

RESULT 7  
B0954456 936 bp mRNA linear EST 21-AUG-2002  
LOCUS B0954456  
DEFINITION AGENCOURT\_8879471 NCI CGAP Co24 Mus musculus cDNA clone  
IMAGE:6476084 5', mRNA sequence.  
ACCESSION B0954456  
VERSION B0954456.1 GI:22369934  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 936)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: gcrabs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL4015 row: n column: 21  
High quality sequence stop: 674.  
Location/Qualifiers

1. 936

source

/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6476084"  
/clone\_1ib="NCI CGAP Co24"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.6 kb. Constructed by life  
Technologies. Note: this is a NCI CGAP library."  
BASE COUNT 207 a 260 c 239 g 226 t 4 others  
ORIGIN

Query Match 17.9%; Score 596.8; DB 14; Length 936;  
Best Local Similarity 85.9%; Pred. No. 2.1e-127;  
Matches 672; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 2331 GCGTGCACCTTGGCTGATGACATTTTGGGGTGACACAGATATGACTCTGATGAGAT 2390  
DB 1 GCCCGTACCTCTTGGACCTTGTGGAGTGCACAGACATGACCTTGTGGAAGAT 60  
QY 2391 CACCTCCCGAGCTGACCTTTCACCTGAGTCCCTGGGAATACCGGAGCTGATCTTCT 2450  
DB 61 CGCTGCCAGTGGAACTTTTACCCAGAGACCTCGGAAATTCGGACATAGCTTCTT 120  
QY 2451 TTATAGTCCATGATGATGACCCAGTCTGCAAGTCTGGGAGATCAACCCATCCGCT 2510  
DB 121 TTATAGTCCATGATGATGATGACCCAGTCTGCAAGTCTGGGAGATCAACCCATCCGCT 180  
QY 2511 CAGGTGCACTCCACAGAAACCTGCTCCCTGGAAGTTTGTGCTGCTGCCAGGAGCTGTCAGA 2570  
DB 181 CAGGTGCACTCCACAGAAACCTGCTCCCTGGAAGTTTGTGCTGCTGCCAGGAGCTGTCAGA 240  
QY 2571 TGGGACCTGTGATGAGCTGCACTTCCACTCTGCTGGGAGAGCGGGTCTGTGCGCCT 2630  
DB 241 TGGGACCTGTGATGAGCTGCACTTCCACTCTGCTGGGAGAGTGTGGCGCTTGTCCACT 300  
QY 2631 CTGCTCAGTGGCTGACTACATGCTATGCTACAGAGCTGTGTGGGTGGGATCCAGAGAC 2690  
DB 301 CTGCTCAGTGGCTGACTACATGCTATGCTACAGAGCTGTGTGGGTGGGATCCAGAGAC 360  
QY 2691 TACTTACGTTGGGAGAACCCAGCTATGCTCTGCTGGGCACTTCTGCTGCGGACAGAG 2750  
DB 361 TACTTACGTTGGGAGAACCCAGCTATGCTCTGCTGGGCACTTCTGCTGCGGACAGAG 420  
QY 2751 AGTCAACATCTGCAAAACCATAGATTTCTGGCTGAAATGGGCACTCTGCAAGGACCTG 2810  
DB 421 AGTCAACATCTGCAAAACCATAGATTTCTGGCTGAAATGGGCACTCTGCAAGGACCTG 480  
QY 2811 TACTGCACTCTGCTACCGCTTGTGACCTGCTACTTTTGGAAAAAAGTAACAAACTAGA 2870  
DB 481 CACTGCACTCTGCTACCGCTTGTGACCTGCTACTTTTGGAAAAAAGTAACAAACTAGA 540  
QY 2871 GTTCAAGTACCTCCAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2930  
DB 541 GTTCAAGTACCTCCAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 2931 TGACAGCTGCGCCATCATGAGAGGAGAGATGAGAGAGCACTCATCTTACACAGCA 2990  
DB 601 AGACAGCTGTGCGCCATCATGAGAGGAGAGATGAGAGATGAGATGAGATGAGATGAGAT 660  
QY 2991 GAA-TCAGTCTTTGGGAAAGATCAATCAATTTTACCTCCAGAGAGACTCTGATGATTTGA 3049  
DB 661 GAAATCCTCTTTGGGAAAGATCAATCAATTTTACCTCCAGAGAGACTCTGATGATTTGA 720





Site\_1: sfil (ggcgcctgcgc); Site\_2: sfil (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCAATTTAGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCCGAGGCGCCGACG-3' (30)Bn-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.2 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 147 a 160 c 156 g 146 t

ORIGIN

Query Match 17.5%; Score 585; DB 12; Length 609;  
Best Local Similarity 99.7%; Pred. No. 1.2e-124;  
Matches 607; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 2424 CTGGGATACCGGACGATCTTTATAGTCCATATATGACCCAGTCTGAG 2483  
DB 1 CTGGGAAATACCGGACGATCTTTATAGTCCATATATGACCCAGTCTGAG 60

OY 2484 TTCTGGAGATCAACACATCCGCTGAGTGCAGTCCAGAAATGTCCTGAG 2543  
DB 61 TTCTGGAGATCAACACATCCGCTGAGTGCAGTCCAGAAATGTCCTGAG 120

OY 2544 TTCTGCTGCTCCAGAGAGTCTGATGAGTCTGATGCTGCAACTTCCACT 2603  
DB 121 TTCTGCTGCTCCAGAGAGTCTGATGAGTCTGATGCTGCAACTTCCACT 180

OY 2604 GTGGAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2663  
DB 181 GTGGAGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

OY 2664 CAGCTGTGTGCTGGATCCAGAGACTTACGTTGAGGAGAGAGAGAGAGAG 2723  
DB 241 CAGCTGTGTGCTGGATCCAGAGACTTACGTTGAGGAGAGAGAGAGAGAG 300

OY 2724 TGTGAGCATTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2783  
DB 301 TGTGAGCATTTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

OY 2784 GAAAGTGGGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2843  
DB 361 GAAAGTGGGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

OY 2844 CTTTGGAAAAAGAACTAGAGTACAGTACAGTACAGTACAGTACAGTACAG 2903  
DB 421 CTTTGGAAAAAGAACTAGAGTACAGTACAGTACAGTACAGTACAGTACAG 480

OY 2904 TCTCAAGAGACTGACCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2963  
DB 481 TCTCAAGAGACTGACCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

OY 2964 AG 3021  
DB 541 AG 600

OY 3022 CCTCAAGA 3030  
DB 601 CCTCAAGA 609

RESULT 10  
LOCUS BF979362 879 bp mRNA linear EST 23-JAN-2001  
DEFINITION 602288452F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4374181 5',  
ACCESSION BF979362 mRNA sequence.  
VERSION BF979362.1 GI:12346577  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 879)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Miklos Palovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihai  
Toshitake and Phero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLNL0037 row: 9 column: 14  
High quality sequence stop: 660.  
Location/Qualifiers  
1..879

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4374181"  
/clone\_1ib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescript (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI; XhoI (gtcgag  
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTTNN-3',  
size-selected for average insert size 2.2 kb and  
normalized to 10^5. This is a primary library enriched  
for full-length clones and constructed using the  
cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NHGRI/NHRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 188 a 261 c 275 g 155 t

ORIGIN

Query Match 16.9%; Score 563.6; DB 12; Length 879;  
Best Local Similarity 97.4%; Pred. No. 1e-119;  
Matches 637; Conservative 0; Mismatches 9; Indels 8; Gaps 6;

OY 4 GAAAG 63  
DB 85 GAAAG 144

OY 64 GAGCGTGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123  
DB 145 GAGCGTGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 204

OY 124 CCCGCGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183  
DB 205 CCCGCGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 263

OY 184 GGACCGGAG 243  
DB 264 GGACCGGAG 323

OY 244 AGCAGGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303  
DB 324 AGCAGGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 383

OY 304 CCTGACCCCGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363  
DB 384 CCTGACCCCGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443

OY 364 ATGAAG 423  
DB 444 ATGAAG 503

OY 424 CGGTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483  
DB 504 CGGTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 561

OY 484 CTGATGACAGTGTGCTGAGTCCAGC---GGAACTGACTTGTGCTCCAGTGGGTTCC 540  
|||||

Db	562	CTGATGACACTGCGTCTGATGATCCAACCCGGGGAAACGTAAGTCTTGCTGTCACAACTGGGTTCCC	621
Qy	541	CGGGCGACCTACATCCGCTCCCTTAAACAGCAGAGATGACA-CCACACTGATGATGACGCCGT	599
Db	622	CGGGCGACCTACATCCGCTCCGACACGAGCAATGACACAGCCACACGATGTCAGCCGT	681
Qy	600	CAACCTGAGCAATGTGGCAGCGTTAACTTGTAATCTACTATTCACAGCATCCAG	653
Db	682	CAA-CTGAAAGCATCTGGCACCAGCGTAAGTCTTGATTAAGTACTACTATTCAGAAATCCAG	734
RESULT 11			
BM790880			
LOCUS	608 bp	mRNA	linear
DEFINITION	k-EST0070769	S21SN0520	Homo sapiens cDNA clone S21SN0520-13-A09 5',
ACCESSION	BM790880		
KEYWORDS	BM790880.1	GI:19139112	
ORIGIN	human.		
ORIGIN	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
	1 (bases 1 to 608)		
	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,		
	Oh,K.T., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and		
	Kim,Y.S.		
TITLE	21C Frontier Korean EST Project 2001		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: kim ys		
	Genome Research Center		
	Korea Research Institute of Bioscience & Biotechnology		
	52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea		
	Tel: +82-42-860-4470		
	Fax: +82-42-860-4409		
	Email: yongsung@mail.kribb.re.kr		
	Plate: 13 row: A column: 09		
	High quality sequence stop: 608.		
FEATURES			
source	Location/Qualifiers		
	1..608		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="S21SN0520-13-A09"		
	/clone_id="S21SN0520"		
	/sex="F"		
	/tissue_type="Stomach"		
	/cell_type="Floating aggregates"		
	/cell_line="SN0520"		
	/lab_host="Top10F"		
	/note="Organ: Stomach; Vector: pTZ18RP1; Site:1; EcoRI,		
	Site:2; Note: The poly (A)+ RNA was dephosphorylated with		
	bacterial alkaline phosphatase (BAP) and then deapped		
	with tabacco acid pyrophosphatase (TAP). The deapped		
	intact mRNA was ligated with DNA-RNA linker including Ecoor		
	I site by treatment of 14 RNA ligase and the first strand		
	cDNA was synthesized from oligo dt-selected mRNA by		
	priming with dt-tailed vector. The cDNA vector was		
	adjusted to have about 60nt. The cDNA vector was		
	clonarlized with E. coli DNA ligase after digestion of		
	Ecoor which site is also included in vector. An RNA strand		
	converted to a DNA strand by Okayama-Berg method. The		
	obtained cDNA vectors were used for transformation of		
	competent cells E. coli Top10F' by electroporation method.		
	The cDNA libraries constructed by this method are		
	full-length enriched cDNA library."		
BASE COUNT	147 a	182 c	126 g
ORIGIN			151 t
			2 others
Query Match	16.8%;	Score 559.2;	DB 14;
Best Local Similarity	98.8%;	Pred. No. 1.1e-118;	
Matches 594;	Conservative	0;	Mismatches 4;
		Indels 3;	Gaps 3;
Qy	1882	TACTATATTTCAGCAGATTTCAGAACCTTGCCATCTCGTGGCCGCCCTTAACACAAATTTCGAAA	1941

[illegible]

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4487127"
/tissue_type="tumor, biopsy sample"
/dev_host="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pcMV-SPO6; Site:1; Sali:
Site:2; NotI: Cloned unidirectionally. Primer: Oligo
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      299 a      298 c      278 g      219 t      2 others
ORIGIN

```

```

Query Match      16.7%  Score 558.4; DB 12; Length 1096;
Best Local Similarity 85.1%; Pred. No. 1.6e-118;
Matches 659; Conservative 0; Mismatches 111; Indels 4; Gaps 3;

```

```

586 CTGATGTACGCGTCACCTGAGCAATCTGGACCGTTAACTTGAATCTACTATCA 645
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 1 CTGATGTATGCCGTACCTAACCTAAGCTGGACCGTTAACTTGAATCTACTATCA 60
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 646 GACTCCAGCATCTTTGAGTTTCTTCAAGATGACAGTCCAGCCCAATGACAT 705
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 61 GACCCAGCATCTTTGAGTTTCTTCAAGATGACAGTCCAGCCCAATGACAT 120
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 706 GACTCCAGTGTATGAGACCAAGCAAGATGAGTAATCCAGTGTGAGCTAAT 765
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 121 GACTCCAGTGTATGAGACCAAGCAAGATGAGTAATCCAGTGTGAGCTAAT 180
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 766 CGAGCAATATGCTCTTATTTGAGAACACAGGCTTCTGATGAGCAAGTACCC 825
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 181 AGAGCAATATGCTCTTATTTGAGAACACAGGCTTCTGATGAGCAAGTACCC 240
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 826 AAGCGTGTCTGTAGAAACATTTGCCATACAGGGGTTGCTTCAAGATGCTTC 885
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 241 AAGCGTGTCTGTAGAAACATTTGCCATACAGGGGTTGCTTCAAGATGCTTC 300
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 886 CCTGCAACCTGACGCTATGACAGCAAGAGGCTCTCTTTTGTCAAACTTTGCCA 945
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 301 CCTGCAACCTGACGCTATGACAGCAAGAGGCTCTCTTTTGTCAAACTTTGCCA 360
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 946 GCCACTCTTATTCATTAAGAGAACTTTTGCACAGTGTGACCTGACAAATATAC 1005
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 361 GCCACTCTTATTCATTAAGAGAACTTTTGCACAGTGTGACCTGACAAATATAC 420
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1006 TCAGGAAGATCTTCTCTGTAAGCTGCGCCAGCTTGCACAGCAAGATATTTTC 1065
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 421 TCAGGAAGATCTTCTCTGTAAGCTGCGCCAGCTTGCACAGCAAGATATTTTC 480
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1066 TACACACAGAGGCTGATGCCAAGAGAGACACAACATGATCAATATGGGCCAAG 1125
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 481 TACACACACAGAGGCTGATGCCAAGAGAGACACAACATGATCAATATGGGCCAATA 540
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1126 CGGAATCTGTAGAGAGACCTTGAAGGAGGACAGTGCCTGCTGTGTGAAG 1185
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 541 CCAAAATCTGCGGTGAGAGACCTGGAAGGGGGGTGAAGTCCGCTCTGTGTGAAG 600
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1186 ACCCATGCTCCAGCTGACAGCCAGGCTTTTCAAAACACAAGACAGCTGCCAGCC 1245
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 601 ACCCATGCTCCAGCTGACAGCCAGGCTTTTCAAAACACAAGACAGCTGCCAGCT 660
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1246 TGCCATATGTTCTTCTACTCCATGAGCTACAGCTACCGGCTCCCTGAGAGCTGAA 1305
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 661 TGCC--ATAGGCTGTACTCCATGAGCTCGGATTTGTACCC--ATGCCAGAGAGGAGCTGAG 717
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Qy 1306 CCGTCTGTGGATTGA--ATACAATGTGTGAGAACAGCTGCCCAACAATGGA 1358
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 718 CTTGCCGGGGGATTGAGTTGCCACCTGTGTGACACCTGACCTCCCAACTGGAA 771
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||

```

RESULT 13

```

AT301140/c
LOCUS      AT301140      564 bp      mRNA      linear      EST 01-FEB-1999
DEFINITION g016905.x1 NCI-CGAP_Lu5 Homo sapiens CDNA clone IMAGE:1908728 3,
            mRNA sequence.
ACCESSION  AT301140
VERSION    AT301140
KEYWORDS   AT301140.1 GI:3960486
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 564)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: c9apbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: M. Bento Soares, Ph.D.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/biopr/image/image.html
            Insert length: 1305      Std Error: 0.00
            Seq primer: -400P from Gibco
            High quality sequence stop: 460.
            Location/Qualifiers
                source
                1..564
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1908728"
                /clone_11b="NCI-CGAP_Lu5"
                /tissue_type="carcinoid"
                /lab_host="DH10B"
                /note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a
                modified polylinker; 1st strand cDNA was prepared from a
                neuroendocrine lung carcinoid, and was then primed with a
                Not I - Oligo(dT) primer. Double-stranded cDNA was ligated
                to Eco RI adaptors (Pharmacia), digested with Not I and
                cloned into the Not I and Eco RI sites of the modified
                pTZ19 vector. Library is normalized. Library was
                constructed by Bento Soares and M. Fatima Bonaudo. "
BASE COUNT 144 a      120 c      157 g      143 t
ORIGIN
Query Match      16.6%  Score 552; DB 9; Length 564;
Best Local Similarity 99.8%; Pred. No. 5.1e-117;
Matches 563; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```



/strain="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5059804"  
/clone\_lib="NCI CGAP Mam2"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"  
BASE COUNT 162 a 227 c 172 g 183 t  
ORIGIN

Query Match 16.2%; Score 539.2; DB 13; Length 744;  
Best Local Similarity 84.5%; Pred. No. 4,6e-114;  
Matches 630; Conservative 0; Mismatches 113; Indels 3; Gaps 2;  
QY 1753 ACCAATGACGTTGCCAAGATCTACTCCATCAATGTCACCAATGTTATGATGAGCGCTGGCC 1812  
DB 1 ACTAATGATGTGCGCAAGATCTACTCAATCAATGTCACAGTGATGGGGGGCTGGCC 60  
QY 1813 TCCTACTGCGCCGCTGCGCTGAGAGCCTGATGTTGGGCTCTCTCGACCTCTTGT 1872  
DB 61 TCCTACTGCGCCGCTGCGCTGAGAGCCTGATGTTGGGCTCTCTCGACCTCTTGT 120  
QY 1873 CCTGCTGTTACTATATGACCGAGATTCAGAACTGCGACCTGCTGCGCCCTTAACACA 1932  
DB 121 CCTGCTGCGCATTCATCAACCGGATTCGTGGACCTGCGACCTGCTGCTTAAACACC 180  
QY 1933 ATTCGAAAGCCGACGACCTTATGTTGTCAGGCGCTGTCCTGCTGTCAGGGACC 1992  
DB 181 ATCTGAAGCTCACACGCTTACGCTGACAGGCGCTGTCCTGCTGTCAGGGACC 239  
QY 1993 AAGAACACAAAGATCCACTCTCTGCTACATGATTCACCTCTCTACGCAACACTCCA 2052  
DB 240 AAGAAATACAAAGATCCACTCTCTGCTACAAAGCTGACCTCTCTCCGAAACACCCCC 299  
QY 2053 ACCAGAGCTTCAACTACACTTTCGCGTTGGCAACACCGTCACTCTGCTGGAAGG 2112  
DB 300 AGCAGAGATTTTCACTACACTTTCGCGTTGGCAACACCGTCTCTGCGAGGGGTG 359  
QY 2113 CCAAGCTTCACTTCCAAAGGTTGAAATACCTTCATCACTTACCTCACTCTCTGGA 2172  
DB 360 CCGAGCTTCACTTCCAAAGGTTGAAATACCTTTCATCACTTACCTCACTCTCTGGA 419  
QY 2173 AACCAAGGTTAGGAAATGCTCTGCTGACCGACATGTCACCTCTCGGATTCCTGAG 2232  
DB 420 AACCAAGGTTAGGAAATGCTCTGCTGACCGACATGTCACCTCTCGGATTCCTGAT 477  
QY 2233 GGTGAGTCAGGTTCTCCAAATCTATACACGCTACGCTGCGCAGGCAAGTCATCCCC 2292  
DB 478 GGTGAGTCAGGTTCTCCAAATCTATCCGTCACAGCTATGTCGCAAGTTGTCATCCCC 537  
QY 2293 CCAAGAGTGACAGGCTACAGGCGGGGTTCTCACAGCCTGTCAGCCTTGTGATCGA 2352  
DB 538 TCTGAGTGATGGGCTACAAAGCGGGGTCTCACAGCCGTCAGCCTTGTGACCGA 597  
QY 2353 CTATATGGGTTGACAAACGATATGATCTGATGGAATCACTCCCAAGCTTTTC 2412  
DB 598 CTGTTGGAGTGTCAAACGACATGACCTTGGAAAGAAATGCTCCCAAGTTGAACCTTTT 657  
QY 2413 CACCTGAGAGTCCTTGGAAATCGGAGCTGATCTTTTATAGTCCAAATGATGAGCC 2472  
DB 658 CACCAAGAGACTCGGGAATCCGACATAGTCTTTTATAGTCCAAATGATGAGCC 717  
QY 2473 CAGTCTGACAGTCTGGGAGATCAAC 2498  
DB 718 CAGTCTGACAGTCTGGAAATCAAC 743

Search completed: March 11, 2003, 08:12:41  
Job time : 3049 secs

